

> O <
O | 0 Intelligence
> O <
FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 09972467-2-589-642-ab037733.res made by jdelaval on Mon Jun 24 Jun 102 6:08:30-

Query sequence being compared: US-09-972-467-2 (589-642)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of US-09-972-467-2 (589-642) with:
File: ab037733.seq

Table with 10 columns: O, F, E, Q, U, N, C, E, S, SCORE, STDEV. Values range from 0 to 54.

PARAMETERS

Similarity matrix Unitary 6 K-tuple 2
Translation Frame 1
Mismatch penalty 1.00 Joining penalty 20
Gap penalty 0.05 Window size 32
Gap size penalty 0
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 13 Median 7 Standard Deviation 19.79
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 10271
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Init. Opt.

Table with 5 columns: Sequence Name, Description, Length, Score, Sig. Frame. Row 1: ab037733, check: 480, 1712, 54, 54, 2.07, 2

The list of other best scores is:

Table with 5 columns: Sequence Name, Description, Length, Score, Sig. Frame. Rows 2-6 with various sequence names and scores.

1. US-09-972-467-2 (589-642)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS SOURCE
Homo sapiens brain cDNA to mRNA, clone\_11b:pb1uescriptII SK plus
Clone:fh11767.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
1 (sites)
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
20181126
JOURNAL MEDLINE
2 (bases 1 to 5139)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@nfo.kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

FEATURES

source
Location/Qualifiers
1..5139
/organism="Homo sapiens"
/db\_xref="taxon:9606"
/clone="fh11767"
/tissue\_type="brain"
/clone\_11b="pb1uescriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="KIAA1312"
/feature="Start codon is not identified."

CDs
CDS
/translation="NSEHFAVVISLCSGMLGTFPSSHGDGVDYIEPLQSGMDSRDEDEEENK
PHITVRSAPORREPSNGRHAQCDPSEKHNRSRKRKTKRPAKMGEPRLNAGDVAALNSG
LATERSAYGKNDNTREKTRHRRTRKFLSTPRPELVVADNRMYSHGENLQHTL
TMSIVASIKKDPISIGNLINVIVNLVHNEEDGSPISFNAQTTLKNCQWQHSKNS
PGIHHDTAVLTRQDLCRAHDKCDTGLAELGTICDPYKSCSISBDSGLSTAFITAH
ELGHVFNMPADHNNKCKEKGKSPDMAPLTFNINPMMSKSRKYTEFEDTGYG
ECLNPESESRPVLPVDPGLIILVNVKQCELTIGPSOVCYPMWOCRLMNCNIVNGVH
KGCQHTHPMADGTEPEBPKKCKKCYGVPEKMDVPTVDSGWSGSPKSRKTCGCGGI
KTATRECNREPRKNGKCYCGRMRKSKCTEELQKRDYFRDQCAGHHDGKHFNIN
LLEPWRVWPKYSGLMKDRCKLFCRVAGNATVAYQLDRYIDTGTGCGDPTNDCVQGLC
RQAGCDHVLNSKARRDCKGCGDSDNSCKTVAQTFMTVHYGYNTVVRIPAGAVINIDVR

OHSEGETDDNYLALSSKGEFLNGNFVTKAKREIRIGNAVVEYSGSEETAVERIN  
 STRDIEOELLVYSGKLYLNDPVRYSFNIPIEDKPOQYPMNSHGWMQACSKPCGGER  
 KRKLVCSRESDDLVYSDORGLNDRPQGHTEPEGCDLRHMVHVASSECSAGCGGYR  
 TLDIYCAKSRLDGKTEKYDDGSSHPKPSNBERKSGECCNAGMRYSGMSSKSCSD  
 GGTORRALICVNRNDVLDSSKCTHOEKYTLQORCEASWQADPWCSTYCGQYQLR  
 KHQWVCQFEDLNDRCMDPEFKPYSQTKOQCEASWQADPWCSTYCGQYQLR  
 AVACILIGTVSVVDDNDCAATRPDTQDELEPESCAPLPPVAKESVSTYPCQMKALD  
 GSWTPCSATGKTRMRYVYSCRDENGSVADESACATLPPVAKESVSTYPCQMKALD  
 CNPTPESERDCQGPFCPLTYWRAEMOECUTCGESGRYKRVVYCVDDNKNKREHVGAR  
 CDVSKRPVDESLSQPCYEVWTTGEMSWPVMEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..

Initial score - 54 Optimized score - 54 Significance = 2.07  
 Residue Identity - 100% Matches - 54 Mismatches - 0  
 Gaps - 0 Conservative Substitutions - 0  
 Translation Frame - 2

NSEHTAVISLCSGMLGTFRRSHDGYFTEPLQSMDEDEEONKPHIITRRSAFQREPTGTGRRHACDTSEHKN  
 10 20 30 40 50 60 70  
 RHSKDKKTRARKNGERINLGDVAALNSGLATAPASAKGNKTDNTRERKPHRRTRKREPLSYPRVEVLYVAD  
 80 90 100 110 120 130 140  
 NRMVSYHGENLQHYILTLMSIVASIIYKDPISIGNLINTVYVNLVIHNEODGPSISFNAQTLTKKFCQWQHSK  
 150 160 170 180 190 200 210  
 NSPGIHHDTAVLLTRDIDICRANHKCQDTLGLAELGTICDPRYRSCSISEDSLSAFLIHNELHGVFMRPHDD  
 220 230 240 250 260 270 280  
 NNKCKEKGYSRPHVMAFTLNFYTNPMWMSKSKRYITFEFLDTGYGCLLNEPESRPPRYVOLLPGILLVYN  
 290 300 310 320 330 340 350 360

KOCELIFGPGSOVCEYMMQCRRLMCCNNVNNGVHKCQKQNTQHPMADGTCESEPKHCKKYGCVCYKPKEMDVPYDGS  
 370 380 390 400 410 420 430  
 10 20 30 40 50 X  
 WGSMSPEFTCSRTGGGKIKTARICRNRPERKNGKCYVGRMRKFKKSCNTEFC  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 WGSMSPEFTCSRTGGGKIKTARICRNRPERKNGKCYVGRMRKFKKSCNTEFC  
 440 450 460 470 480 X 490 500

NINGLIPVNRVPRYSGLTMMKDRCKLRFVAGNATAYVOLLRDRVIDGTPCGQDPTNDICVQGLCRQAGGDHYLN  
 510 520 530 540 550 560 570  
 SKARRDKCGVGGDNSSCKTVAGTFENYVGYNTVVRIPRACATNIDVROHSFSGSETDDNYLALSSKGEFL  
 580 590 600 610 620 630 640  
 LINGNFVTKAKREIRIGNAVVEYSGSEETAVERINSTRIEQLDLYLVSGLKLNPDVRYSPNPIEDKPOQ  
 650 660 670 680 690 700 710  
 FYWNSHGFMQACSKPCGGERKRLVCTRESDDQLTVSDQRCDRLPQPHGITPERCGTDCDLRHWNVASRSECSAQ  
 730 740 750 760 770 780 790  
 CGLGYRFLTDIYCAKSRLDGKTEKYDDGSSHPKPSNBERKSGECCNAGMRYSAMTEGSKSCSDGQGTORRRA  
 800 810 820 830 840 850 860  
 ICVNRNDVLDSSKCTHOEKYTLQORCEASWQADPWCSTYCGQYQLR  
 870 880 890 900 910 920 930  
 TKPYSMQCQDPEASWQADPWCSTYCGQYQLRRAVKCTIIGTVMSVVDNDNNAATRPPTDQDGLPESGH  
 940 950 960 970 980 990 1000

PPPAEETRSTYSAPTQWRGSGWTPCSATCGKTRMYVSCRDENGSVADESACATLPPVAKESVYPR  
 1010 1020 1030 1040 1050 1060 1070 1080  
 CGQWKAIDMSSCVTCGGGRATRQWVYVSDHYIDRSQCDDDYIPEPDQDMSMPCPQRPRDSSGLAQHPRQ  
 1090 1100 1110 1120 1130 1140 1150  
 NEDYRPRASPSRTHVLAGGNQMRGTGPMWAGSSCTAGSGQRVYVODENGYTANDCOVERIKRPEDEORCESGP  
 1160 1170 1180 1190 1200 1210 1220  
 CPQWAAQNMNCEFTKICGGGIRTRILYVQCSNBERPDLSCETLIDKPPRDEOCNTHACHDAMSTGPMSSCS  
 1230 1240 1250 1260 1270 1280 1290  
 VSCGRGHKQNNVYMAKDDSHLESYDCKHLAKRPHGHRKCRGRCRPMKAGAMWQSCVSGRCVQQRHVQCI  
 1300 1310 1320 1330 1340 1350 1360  
 GFPIAKIRETECNPTPESERDCQGPFCPLTYWRAEMOECUTCGESGRYKRVVYCVDDNKNKREHVGARCDYS  
 1370 1380 1390 1400 1410 1420 1430 1440  
 KRPVDESLSQPCYEVWTTGEMSWPVMELXRPSSAQWPESGRGMSGAMWETXGYRPPDLPLXHYVCFL  
 1450 1460 1470 1480 1490 1500 1510  
 TSPQLSOARITPCMQSGRDVDDGLYVPSGISHVITHDIVLAKRNRYYRGGCKLIGCEVEPDSVYXENIYVPGI  
 1520 1530 1540 1550 1560 1570 1580  
 QOGEATVYNXHMILCEGSRISFTPAVAMQCGIALPDLIFQERLKITWTSMDPLTXKWKGLFLYLKHKCGQT  
 1590 1600 1610 1620 1630 1640 1650  
 NPHEBEDVLCLENKQQRKRIIFVYNRISSTIFSYIFLCTMLIYTLFNGKRPVYI  
 1660 1670 1680 1690 1700 1710

2. US-09-972-467-2 (589-642)  
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139  
 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000  
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.  
 ACCESSION AB037733  
 VERSION AB037733.1 GI:7242978  
 KEYWORDS  
 SOURCE Homo sapiens brain cDNA to mRNA, clone\_lib:pbJuescriptII SK plus  
 ORGANISM Homo sapiens  
 Clone:fh1167.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (sites)  
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.  
 Prediction of the coding sequences of unidentified human genes.  
 XVI. The complete sequences of 150 new cDNA clones from brain which  
 code for large proteins in vitro  
 DNA Res. 7 (1), 65-73 (2000)  
 JOURNAL 20181126  
 MEDLINE 2 (bases 1 to 5139)  
 REFERENCE Ohara,O., Nagase,T. and Kikuno,R.  
 Direct Submission  
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,  
 Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba  
 292-0812, Japan (E-mail:cdna@foc.kazusa.or.jp.  
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,  
 Fax:+81-438-52-3914)  
 FEATURES  
 source  
 1..5139  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="fh11767"  
 /tissue\_type="Brain"  
 /clone\_lib="pbJuescriptII SK plus"  
 1..4417  
 /gene="KIAA1312"  
 <1..4417  
 /gene="KIAA1312"  
 /note="Start codon is not identified."

```

/codon_start=2
/product="K1A1J12 protein"
/protein_id="BAA92550.1"
/db_xref="GI:7242979"
/translaton="NSEHVAVSLGSGMLGTPRRSHDGYETLEPLOSMBDEDEEONK
PHIYRRAKPOREPESTGRACDTPTEHKNRHSKDKKTRRKRKGEINILAGDVAALNSG
LATEAFSAIGNKTTDTRERKRRRFRKFLSTYREVEVLVANANRVSYHNEVLQHYIL
TLMSTVASTYKDPISIGNLNIIVNLIVHNMDSPPISFNQOTLTKRMCQOHKSN
PGIHHHDVAVLLTODICRAHDKCDLGLAEITCDPRCSI SEDSGLSTAFPIAH
ELGHVHNMEDHNNCKEKEGVSPQHVAPALNTFNTMMSKSKRSRKYITTEFLDRG
ECLLNEPESRYPYLPYOLPGLILYNYKCEELRFGSGOYCPMOCRMLMNNVAVH
KGRTRQHPWALDGTGCEPEKHKYGFCEPKENDVYVTDMSGMSPFSTCGTGGLI
KTAIRECNREPKNGKCYVGRKRMFKSCNTEPCQKRFDEEDCAHDFHNFING
LPLNPMWPKYSKARILMKDRCKLFCRVAGNTAYQLDRDITDTPGODTNDICVGLC
ROAGDHVAVNSKARIDKCGVCGDNSSCTVAGTNYGNYVRIIPAGTINIDVR
OHSEGETDDNDVYLAASSKGEFLNNGVYVYMARBELRIGNAVVEYSSEDAVBRN
SNDRLEOELLQVLYAVGKLYNPDVRSFNPIREKRPQYRMSHAPMCAKSCPCGGER
KRLVCTRESDDLVSDQCDLPOGHITTEPCGDILRHVAVARSISACQGLYR
TLDIYCAKYSLRIDKTEKVDGFCSSHPKSRKESGECNNGWRYSAMTECSKSD
GGTORRRAICVMTNVDLSDSKTQHOEKTIDRCSELPQPMKSGDMEELVTCGKH
KRWVWCQGEEDRLNDRMCDPEPTKPTSMOTCOQPCASQWAGPWCOSVTCQOGLR
AVKCIIGYVAVVDNDICMAATRPNDTODCELPSGHPPAARPEMRYSAPRQDWR
GSMTPCSATCGKSTMRKYSCSDENGSAVDAESACKTILRRPAKESCVTPOGKALD
WSSCVYTCGQGRATQVMCVNIDSHVIDRSECDQYIPEFDGCSMSPCQPTPDSGL
AQPPEQMEDYRSPASPSKTHVLDGNQMTGTGAGCSSTCAGSQRVYVQDENGTYL
ANDCVERIKPDEORACESGPCQMAVYGMWEGCTKLCGGGIRTLVVCORRSMGERPDL
SEIILDKPREDCNMTACPHDAAMSTGPMSCSVSCSGRKHORVYVYCMARKSHLES
DYCKHLAKPHGRKGRGKPKKAGAMGSCVSGRQVQOERHNGCOIJGFKLAEFE
CNPYTRPESRDCGPRCRLYTWRAEMQDECTKRTGEGSRIRKRYVCVDNKNRHHGAR
CDVSKRPDRSCSLQPCBYVITGEMSEVSPWEL"
BASE COUNT      1471 a      1112 c      1362 g      1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
Initial Score = 7 Optimized Score = 12 Significance = -0.30
Residue Identity = 24% Matches = 17 Mismatches = 37
Gaps = 16 Conservative Substitutions = 4
Translation Frame= 4

```

```

MITGLLPLKRVXISYVHRKIYEKMWLEILRIILKIIFFLXSLMLNHNSSGSKGKGVPCDLOGCFKTKNPSHP
10 20 30 40 50 60 70
FOYRKSIIILVQIFLSYKXIRSGNARPHFCMATAGVKEILLRSHSMCXFYVVAISPCCIPGTTTFYXHLSSG
80 90 100 110 120 130 140
TSDHPSLQPLXXFLLYANRTRXKIFTECPREBGTXYRPSLTSPLCIGCITLAWESXGLVRRKHTVYSGSGGXYPQ
150 160 170 180 190 200 210
VSHIPLALIPRLIPRLPSLGHGXADDDGYSODGTSDBHSPVIOITYSOGCKLQLSRSTGRILLTSQARAPCTFLISS
220 230 240 250 260 270 280
THTTLRYLEPSPQVLYVHSHSSALQVYKRGHRGPMQSRSDSGLYGGLHVSVALILCVPIXQPCDLCCTPRPQD
290 300 310 320 330 340 350 360
TENHLQARAFHGLHPRRHFLLCPCGLARCLQXSLSKWLPSPFAMQXTFRCCLPRPQTEHEHLEQGPLVHNAASCG
370 380 390 400 410 420 430
QACVLYHCSRSBSGGLSRIQSOLKSGNRSPLRDXQQTSTLYLMPRRHSHVNSRQPLRANHXQGPRPSQALCSSGFLLS
440 450 460 470 480 490 500
TQSFVAVVYRFSSXHTTTRRMDRPAQVLEHNAFOGVLYLHFWPSTWVRLGLALRGRXSSFWKGCXAKRPSGLVXG
510 520 530 540 550 560 570 580
HGDMEQMSVSGIXSMWSHSLRSTTWSLXLNHTCRVALRPPQVYLFQELQSKARHCPRQGVTEHSSPFAIQLGRV
590 600 610 620 630 640
AVALSATEPEPSSKQQLYLRVPRPVAEHLGVODRNRHNVLGAIDYLLVLSGAAGGGXHDGNSQSWVSGVLY
650 660 670 680 690 700 710 720
AALQSLSTPDIYVMMHPTALSWYRPOVLTINRPOGRACODAHSGCXQVCIQVGLVSGSHLISLFRSSPNX
730 740 750 760 770 780 790

```

```

HQTWRCLCPFPQVYTKHSDQSPDFHCGQGNLSHLXMTVTSXCVNLHSSSTSFRLTQIALMLLVWPSPQLFERS
800 810 820 830 840 850 860
VQAEKXQRPVLDLSPHEPFSLLDGLWLDLQKRPSTSPVPPSSLLYTLAQXMSNVRXKRPNHALLSLLATGHLRS
870 880 890 900 910 920 930
QSVPOGQVWPCPQWGRSKNHLXSETVXSDSLVQTSFRFRSPMQLHACHGRYLFQXNCGXLSIGLIEKERT
940 950 960 970 980 990 1000
SGLYNFPDRTKXSKSSCSNRVVELLSTPAVSDPLYSTFAFMRKISLAIYVTTKPEFSKNSPDLIDKAKXLS
1010 1020 1030 1040 1050 1060 1070 1080
SMSPFETGSR-----GGGKIT--AIRG---NR-----EPKNGKYYCVGR--MKTKSCNTEPC
10 20 30 40 50 X
SVSEPKLCCRTSILVAPAPAIRFTVLPKXCTVLTNVPATVTLHEELSPPTPHLSLRFEFKTYSHRACRQRPWT
1090 1100 1110 1120 1130 1140 1150
QISFVWPGVSTILSRKXAAVPRATLQNNLHRSFTIRIPRYLSTQRTLSRPLMKCFPSKKAHCSRRS
1160 1170 1180 1190 1200 1210 1220
RCLRHGVSVDLDFLREPTQYFPFFGSGLLHSRMAVLMRPVHVEQVPKGLDPRDPSVGTGSIISIGTON
1230 1240 1250 1260 1270 1280 1290
PYLQCFPSGSHVSPSAQVXVYQPLCTPLTLIDHQRSLCITYGHXHEGPKTINSCLFTLLKRMPSQWTKGKX
1300 1310 1320 1330 1340 1350 1360
GLDQSGSLSKHSEXPVSKNSVIFYRLLFHDHIGLVYKFKFSVGAATCWGLTLPSSLHLLSSXGMLNTWPPSSMAI
1370 1380 1390 1400 1410 1420 1430 1440
VNAVLPISLLEBOLLGSGQWMPSSARPKVSHLSXALDSCLVKRTAVSXMTPEGLFEECGHMOKFVWVYA
1450 1460 1470 1480 1490 1500 1510
LKDMEGPPSCSLXITTKFTTILLIKRPIILGSLXIEATIDIKYKIKXCRSPWYETLISATTKTSTNRGXDK
1520 1530 1540 1550 1560 1570 1580
RFVLLMVLVLSVLYLPAEMASVAKRPFNAATSPARLILSPHRLALVFLSILCLFECSEVSHACLPVBS
1590 1600 1610 1620 1630 1640 1650
SLMGALRLMMWGLRCCSSSSSIDCGSGSIIKXKSPSXDNRNVPISPEQRIKMTAVCSE
1660 1670 1680 1690 1700 1710

```

3. US-09-972-467-2 (589-642)  
 AB037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139  
 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000  
 DEFINITION Homo sapiens mRNA for K1A1J12 protein, partial cds.  
 ACCESSION AB037733.1 GI:7242978  
 VERSION AB037733.1  
 KEYWORDS  
 SOURCE Homo sapiens brain cDNA to mRNA, clone\_h1b:pb1uescriptII SK plus  
 clone:fh11767.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (sites)  
 AUTHOR Kikuno, R., Ishikawa, K. I., Hirose, M. and Ohara, O.  
 TITLE Prediction of the coding sequences of unidentified human genes.  
 XVII. The complete sequences of 150 new cDNA clones from brain which  
 code for large proteins in vitro  
 JOURNAL DNA Res. 7 (1), 65-73 (2000)  
 MEDLINE 20181126  
 REFERENCE 2 (bases 1 to 5139)  
 AUTHOR Ohara, O., Nagase, T. and Kikuno, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,  
 Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba

292-0812, Japan (E-mail:cdna1nfc@kazusa.or.jp, URL: http://www.kazusa.or.jp/nudge/, Tel: +81-438-52-3913, Fax: +81-438-52-3914)

FEATURES Source Location/Qualifiers

1..5139 /organism="Homo sapiens" /db\_xref="taxon:9606" /clone="fh11767" /issue\_type="brain" /clone\_lib="pbluescriptII SK plus" 1..4417 /gene="KIAA1312" <1..4417 /gene="KIAA1312" /note="Start codon is not identified."/ /codon\_start=2 /product="KIAA1312 protein" /protein\_id="BAA92550.1" /db\_xref="GI:7242979" /translacion="NSEHTAVISLCSGMLGTPRSHDGYFIEPLQSMDEQDEEEDENKPHIYRSAPQREPSGTRHACDISEHKNRHSKDKKTRAKMGERTINLAGDVAALNSGLATEPFAVGNKQDNTREKRRHRTKRLSPRPVAVLVYADNPMYSYSEMNQHYTLTMSIYAIKDSITGNLIVNLIVHNEODGSIENRQITLKNFCQWHSKNSPGIHDHVAULTKQDTCRAHDKCDTLGLAELGTLCDPYRSCSISEDSGLSTAFIAH ELGHVFMPIHDNNKCKEVEKSPQHWAPLTFEITPMWMSCKRYITFEIDTGYG ECLLNBESEPRYPVLPVLPGLLYNVNKCCLLITGPGSQVPMYMCQKRLMNNVNGVH KGRTOHPMADGTECEPGRKCKYGFCEPKEMDVPVYDGGMSMSPGTCSPRTGGGI KTLARCNRRPKMKGKCYVGRMKFKSCMPEKQIQRPEDEOCANHEPRKPKINGI LLRWVWPKYSGILMKDRCKLRVAGNAYYQIDRVTYDPCGDDPNTDICEGLC RQAGDVAHNSKARDKGVGCGDSSCKTVAAGFENVHGIYVRIEPAINIDVR QHSFSGTDDNNTALSSKGEFLNNGFVYVTAARIRIGNAVVESSGSETVERIN STDRIBELDLQVLSVGLKLYNPVDRYSFNIPIDKPKQVNSHSPWQASCKPQGER KRLLVCTRESPLDSDQCDRLPQGHITPEGCTDCLMWHVAHSESAOGLGR TLDYCAKYSRDLQGTETKVDGFECSHPRKSNREKSGECCNFGMRYSAATEKSKSGD GGRORRALICGMPNPNVDLDSKCHNOEKVITQCSERPQOMKSGMSEGLVYCGGR KHQVWCFGEEDRLNDKACDEPRTPTSMQTCQDEBACSMQAGMGGCSTVCGGQYLR AVKCIIGTYSVVDNDNCNATRPRTDQDELBPCHPAPAFTRRSTYSAPRTWRF GSNPSCATCGKGRMYVSCRDNESVADEASACATLPRVAEBCSVPCGQMKALD WSSCVTCGGQRATQVAVYSDHVADRSECCQDIPETDCCSMSPQRPDPDGL AQHDFQNDYPRASPSRTHVLGQWQPMWAGSITCAGGSSORRYYVCOENEGYT ANDCVERIKDEDEACBSPGCPQWAVGNMKECKIKGCGGIRFPLVYVCOERNGRRPDL SCELIDRPPREOCNTACRHDAAWSTGPMSSCSVSGCRHKQNYCMAKDSHLES DYCKHLAKPHGRKCRGRCRQKWAAMSOCSVSGGVQQRHVGCQIGTKLAREE CNPRTRESEHDCGPRCPLTYMRAEEMOECTKTCGEGSRYRKYVVCVDNKNVHGAR CDVSKRPVDRSCSLDPCFEVYVITGEMSEVSPWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803

Initial Score = 6 Optimized Score = 10 Significance = -0.35 Residue Identity = 22% Matches = 12 Mismatches = 36 Gaps = 6 Conservative Substitutions = 0 Translation Frame = 6

DYWFETIKKGLDKHQAQEDIXKNGTSDTSDIENYFLPKLFLVQAQOHIFMKGVCLLRTEFMXNQKRPQPF 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280

LF---HSCF---HPSHTPSATLRPWLKSKRMLOFRPGRYLXPRFSCDDIAGLQYTAFTVNRPLDANVTACRMYL 220 230 240 250 260 270 280

VFWIHTNHLVADGFAAGLGAELPFLCPRPSVEGTPMALVAERLRSGVWVALGSGEYVCSDDLTAHMBLLY 290 300 310 320 330 340 350 ASANGHALLRPSASFRRGASSSSALSVPMWLSOVLVYITFKMASITGHAVNISLEPMPSTTRRTRAPRASTP 360 370 380 390 400 410 420 CSVVWVSMCVTLTLTGFRFKNFTAOIWKRPFTVGPLDLDHOSCSYASTOTLSALSPVALISPLRTRRAGFTGSLI 430 440 450 460 470 480 490 RFSYLHTVCGVSTFILLTWNTPDLGASNTGTCSPGASSPVLSTEMHGAAGARCTVILLIEGVLXAYTV 500 510 520 530 540 550 560 570 WGPLRAWXHGTVLWSEMDIILVTLTRPDHVVTVVDTNHLRPGCEPLTRGHRARAVQGLDLPRTGCRHFFLCH 580 590 600 610 620 630 640 WNRQSGTGLVSHRALLISADVYSHQCTFPFSGKANGRRKSPDLGSMCVYVCAAPRRRSGSMGNTXKXFTVLG 650 660 670 680 690 700 710 ISMSSCGITVIVYHKHISPNDAFHCSXLVLSLSTSDTALSPTGTRLDPGCTFRLLTSLHRQWGLRVTNHSIIXS 720 730 740 750 760 770 780 IFPKLTGDLAVLMPFSTGDDALXIPVSRFLMTRBELTARPLNGNRLMCAFAVLOYIISGIDTJNSPDLGPTVT 790 800 810 820 830 840 850 AFXYTSPGRJAPATRYTFPFAFFTVAWFGMAAKTIIINLSDRIPARIGTVDVOCAYQATLGTTFPBAQN 860 870 880 890 900 910 920 930 MPPQVYVCTTGTFGFSVNSRHLGQPIASLIRNSKLIIRPGANKRSPFPPLAGFTACLPMWPTVAVKLLRPIFMNI 940 950 960 970 980 990 1000

ERLAVIGVQLSHRONIKOKFELNLAIQXVNSFYGLGPTVLYHSIPNANPFEGCDNKKVSIKQETFTAXXS 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 XVVYVYCFEYKTEVLPKPHINICSTSMNSDHSIVTITMYCICKSCGCEFAKRIITATPTPTISPGFYXNMIASSLP 1150 1160 1170 1180 1190 1200 1210 AKALIDTDICVLAATRSSIYHSVSKLIVGCPCHSAEOLADVLDHONSTVERDPAHIGKQTVDMKMLPVKYSTL 1220 1230 1240 1250 1260 1270 1280 1290 FIVSVDLLEFVLRVAGLKFHSTSYVTFSTIIFMFWVALSNGCFDAPSPCSGASGKPTSPGCSICHGDIHF 1300 1310 1320 1330 1340 1350 1360 FGNFKSLIAYLSRLALRPIPGCVLSPALFVYSIDVIAPRPSALHHIWAHLRFTWKSNOQTTLFTHVVKDMQL 1370 1380 1390 1400 1410 1420 1430 DRQVSGSFRPVRVQTLAITSVXKLSIDIEFTTLKXPHRQVGVQVQWMSHMDLGTLSNFEFFVYVIMBRVYKMA 1440 1450 1460 1470 1480 1490 1500 QMGDRKSTQSTIETFTNRJASIGITNGSQSPQAXGTFEVVSSADLISCEBENSIMMDSTWYLRMLPLAKVF 1510 1520 1530 1540 1550 1560 1570 XCSLSTIKRYGRIPLFIIMNHNXVNNNINXISNTWVFLDRGYNKHXSNXNLYLKVYFSAVXNHSYVCHHQDFYKP 1580 1590 1600 1610 1620 1630 1640 WIGKATPCPSVGLPFCVNBRLITSRKCLCCXAAVXCCYVTSQVNPSPSPSGGFLIVETVPIEYVXGVTM 1650 1660 1670 1680 1690 1700 1710 PSCXGLSLGGAAPINDVGFVLPFLFIFLIRHLXWPKIIPIMRPECQAHSHXABADDGRVIG

4. US-09-972-467-2 (589-642) TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds. ACCESSION AB037733 VERSION AB037733.1 GI:7242978 KEYWORDS Homo sapiens brain cDNA to mRNA, clone\_lib:pbluescriptII SK plus clone:fh11767.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo; 1 (sites)  
 REFERENCE Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,A.M. and Ohara,O. Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro  
 JOURNAL DNA Res. 7 (1), 65-73 (2000)  
 MEDLINE 20181126  
 REFERENCE 2 (Phases 1 to 5139)  
 AUTHORS Ohara,O., Nagase,T. and Kikuno,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:odnainfo@kazusa.or.jp) URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914

FEATURES  
 source Location/Qualifiers  
 1..5139  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="fh11767"  
 /issue\_type="Brain"  
 /clone\_lib="PbluescriptII SK plus"  
 1..4417  
 /gene="K1A1A1312"  
 <1..4417  
 /gene="K1A1A1312"  
 /note="Start codon is not identified."  
 /codon\_start=2  
 /product="K1A1A1312 protein"  
 /protein\_id="BA092550.1"  
 /db\_xref="GI:7242979"  
 /translation="NSEHVAVISLCSGMILGTRRSHDGYETIPLQSDNDEDEEONKPHIVRRAAPORRSPRGRACPTSEKRNHSDKDKTKARRKGERNVAAGDAALNSGLATEASAYGNKTDNTRKTRKRFSLYRFVAVANRWLSTYHGMLOHYILTLMSIVASLYKDPISGILNIVLIVLHNDQDPSISFNQTLKRFCCMQHSKNSPGIHHDDAVLLTRDIDICRAHDKCDLGLAELGTLCDPYRSCSISEDSLSAFTIAH ELGHVYNEHDNNCKEKGKSPQHMVAPLNFPTNPMWMSCKRKYITLEDLDYGECLLNEPESRPPYLPOLBGLIYNNKOEELIFGGSGOCPAPMOCPRMCCNNGVHKGRROHTPMWAGTTEGEBGKHKYGRCEKMDVYVYDTSNMSNRPFGTCSGTGGI KVALRECAINRPERKNGSKYCYGRMKFKSCNTEYQLRKDEDEDECAHDKGHEFNIG LLENVAVPKYSGILMKDKKLCFRVAGNTAVYQLRDIYDTPGCGDINDICVQGLC RQAGCDHVLNYSKARBDKCGGDNSSCKTVAGFTNFTVYGVNIVRIIPAGAFNIDVR OHSFSGETDDNYLALSSSKGFEFLNNGNFTVMARKEIRIGNAVVEYSGSETAVERIN SFDRIEQLLELLOLVYKGLYNDVYRSENIPIEDKPOQRYMNSHGMQACSKRPOGGER KRKIVGTRSDOLIVSDDORCDLROGHTTEPCGDDIRHNVAVRSRSCSAGGIGYR TLDIYAKYSRLDGTKEYDDGFCSSHFRPNSREKCSGICNTGMRYSAMWECSSKSD GGQRKRAICVNTKNDVLDLSDSKCTHDEKVTIQRCSEFPQPMKSGDWSCELVTCGRH KHRQVWCQFGEVLDNDKCDPPTKPTSMQTCOAPRCSAQAGPMQSCSVCQGGQLR AVKCIIGTYMSVVDNDCAATAPTDQCELPSPCHPAPAEFRRTSYTSAFRTQWR GSWPTSATCGGKTRMYRYSKDENGSVADSEACALPRLPVAKEKCSVTYPCGOMKALD MSSCSYTCGGGRATQVMGVNYSOHVIDRSEGDQYIETPDQCSMSRCPQRTPSGL AQHPNEDYRPRRASAQSTFHYLGNOMRTGQWGCSSSTCAGSGSRRYVYVCDERFPII ANDCYERIKRFDQKCESGFCQMAIYGNWDETKLGGGIRTRLRVYCDRNSWNERPDL SCEILDKPRDQKCHVACPHDAAMSTGSSVSSCSGNGHRQVYVCDRNSWNERPDL DYCKLHAKPHGHRCKGRCRPMKAGAMSGSCSVCSSGNGHRQVYVCDRNSWNERPDL CNPRTYRPESEDDCGPRCPDLYTWMRAEMOECTKCTCGESRYYRKYVVCVDDNKNEVHGAR CVYSKRPVYRSCSLOPCEYVITGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t  
 ORIGIN  
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..  
 Initial Score = 6 Optimized Score = 9 Significance = -0.35  
 Residue Identity = 16 Matches = 11 Mismatches = 43  
 Gaps = 12 Conservative Substitutions = 0  
 Translation Frame= 5

XLLLVYHXKGSXALCTGRYMKRWLYRFGYRKLFFSSSEAVCXGTGTHLVERGILFAPYVNLKPKISAPI  
 10 20 30 40 50 60 70

FRLGNLTYXSRFLASLEKGLVTLIDHSAMPQLEKRRKPCSLHTAASLSQLPSLLVLSALQYFLILNCLOA  
 80 90 100 110 120 130 140  
 PHSILKYCSPCNFFPMPGHOGEQNVLDKLRDHLHLHSFRAYKKEFYLXGAEDLSENTQPTVANRAGSLK  
 150 160 170 180 190 200  
 SPFLHSLHSLSCHQAMAEIIMVYVTRTPVPLTILKSRHNRVAVNSFHHGPRPAACSHSYPNAPRRCCHP  
 220 230 240 250 260 270 280  
 HHPGCGTWSLRRRSCILAIPLPESKGRDPTVIGSKRVPYVWCMQCTRSMLWLCVFRSDSPHASAVRIGHR  
 290 300 310 320 330 340 350 360  
 OSTDSKROLSIMWIFLGTFCANVAXGAYNSHNFLLNGFHLILCCKNHFYVYALDHRKONTSSSGOYSMRVD  
 370 380 390 400 410 420 430  
 KHVCTYANDREYVQEFHSSNLETFVHRWADRPVPLFLCLNHTAAXTLPSCKRPTEDKGRHRLHAYQLFSP  
 440 450 460 470 480 490 500  
 HSHLRCTIHFNPRDIOHAAGITRQHRVYNNMLPRGQFSTGFHRAHGCGMWRSGDDSPHFGRAELSHCIGSPFBG  
 510 520 530 540 550 560 570  
 MYTWSNPGQFLGNPENTHSDSRGHCXYHTSLAGLRYLHRSQSKSSPSPSIARHVSQNLPLRPLVYAGX  
 580 590 600 610 620 630 640  
 HRHSRPOQSHSLGSRYSXSHKMLSMGSKTQATGFLVHCQCFSESPGQLEDMMVYHSPGYQLVYL  
 650 660 670 680 690 700 710 720  
 LHSRCHLPLTYLTSQXCSLILLAGLRYVHXHCTVPRDPRRHHIIRADKSAKMLMSGGHTFYHLIDHQD  
 730 740 750 760 770 780 790  
 TRFGAYVALFHRXPSTLTLQISTVDKGTNCTSEWXLSDVLCICCPHYVHNFYXHKRPFSSGSHRSFLNIQ  
 800 810 820 830 840 850 860  
 SRONSASHCTYPLSTFNHCLWYDQCCCKNHNPOSSSHRACIYHNSKCPDMGCSNBTGHIYHSCQNAHATSGH  
 870 880 890 900 910 920 930  
 SLVHRVQXCVQAGADRIDPOKXADQDLRQKQVYVYPLGIVYCMAMANDCSSKTAEVYLDLEYYKXNSVH  
 940 950 960 970 980 990 1000  
 RGTTRPPTKREKVALVLAQODDLSXFFELRQSPNHCPLRQHSCEPFRPWLXOQSFILARHILYCLIKLSSGR  
 1010 1020 1030 1040 1050 1060 1070 1080  
 LPRLRCAANAHOYWHQDLDFRGPQYCNHNVLYXMFPLRFLCMKMYRHHKPHIYLSGLISLKHRIQLAGKPGH  
 1090 1100 1110 1120 1130 1140 1150  
 RYHLCPGKHEFLSLCLEADSRILSLPLRCTTCTGPPSEFFHCIXGPRSAHWEDARXGXNAVSROSEHTVHLSL  
 1160 1170 1180 1190 1200 1210 1220 X  
 GSMSPGTCSTRT-----CGGGIKTAIRECNRPE-----PKNGKRYCVGRMKRKCSCNTEPSC  
 1230 1240 1250 1260 1270 1280 1290  
 ASXKMDAPRCRXYISFYVLSIFHNFELVLCSTLEMLFXCPRLMFMWRSRFORSDRRIHLSRGNPFLMEHKI  
 1300 1310 1320 1330 1340 1350 1360  
 HHCSAQARTPSRHPVCAEAGSGLVXHXKXCTRAVCTASVWGTPENLVOKSIIHYVSRCKGCLAVGQAKGR  
 1370 1380 1390 1400 1410 1420 1430 1440  
 VNIQVAXANTRNNOCKIQXYTFDYTLTTSRGMCRSSVLEPRYHAGDSXLLLLYLITCCCHNEACXTHGPARHRS  
 1450 1460 1470 1480 1490 1500 1510  
 XKLYSIHVLDHXXNSFRDHRKWPVQDGLRYHTCRSLCYRPLYLRBOQVHNDGFHLDCCSSNAATGKSFMLTSH  
 1520 1530 1540 1550 1560 1570 1580  
 XKITWANHRYHNSQXQXOXYLNFQYGLYRXRLQPLFKLYSVYEGFLHGMKDFCCIPRPLLDQYVDRIKN  
 1590 1600 1610 1620 1630 1640 1650  
 VLSFGSSFLCCPSYVHNKMKPILLISRLMLLRHROPGXSPFLPIFLWLFSSCLVCAVYCVLRCNHNAFLDRA  
 LSGGKCAIKXCGVCEVPLDLHLVHPXTVVVQXNNPHEHTGMCPRALFSLRGXKRPCAR

5. US-09-972-467-2 (589-642)  
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000  
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.  
ACCESSION AB037733  
VERSION AB037733.1 GI:7242978  
KEYWORDS  
SOURCE Homo sapiens brain cDNA to mRNA, clone\_lib:pbuescriptII SK plus  
clone:fh11767.  
ORGANISM Homo sapiens  
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
Bakaliya; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
EPERENCE Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.  
AUTHORS 1 (sites)  
TITLE Prediction of the coding sequences of unidentified human genes.  
XVI. The complete sequences of 150 new cDNA clones from brain which  
code for large proteins in vitro  
JOURNAL DNA Res. 7 (1), 65-73 (2000)  
MEDLINE 20181126  
REFERENCE Ohara,O., Nagase,T. and Kikuno,R.  
AUTHORS Direct Submission  
TITLE  
JOURNAL

FEATURES

source location:Qualifiers  
1..5139  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="fh11767"  
/issue\_type="brain"  
/clone\_lib="pbuescriptII SK plus"  
1..4417  
/gene="KIAA1312"  
/gene="KIAA1312"  
/note="Start codon is not identified."  
/codon\_start=2  
/product="KIAA1312 protein"  
/protein\_id="BA92550.1"  
/db\_xref="GI:7242979"  
/translation="NSEHTAVLSLCSGMLGTFKSHDGDYFIEPLQSMDEDEDEBONK  
PHIITRRSAQREPTSTGRHACDITSEHKNRHSKDKKTRAKRKGERRINLAGDVAALNSG  
LATEAFSAVGNKTDNTRKTRKTRKPLSYFYFVLLVADNRMRVSYGEMLDQHLIL  
TMSIVASITKDSIDGNLNIYVIVLIVHNEGDGSISEFNQDTLTKNCCOMHSKNS  
PGGHHNDPAVLLTRDICIARHDKDPTLGLAELGTIDDPYRSCSISPDGLSTAFPTAH  
EDGHEVMPHDDNNKCKEBSVSPQHVMAPTLFTNPMWKSCKRKYTTETLDITGCG  
ECLLNPEBSPYPLVQLPGLILVNNKQCELEIFGPGSYVCPYIMQCRKMLNMMNVNH  
KCGRTQHTPWARDTECEPGRKNGKYGFCVPEKEMVNPVTDKDSWGSWSPFGCSKRGGI  
KTAIREECNRPNKGGKCYVGRMRKFKSECLCOKSDRPEOCANHDGHNFNING  
LLENVWVPKYSGILLMDRCKLFCVAVGANAYVYLDGPDYVDDGPPVADNDICVGLC  
RQAGCDHVLNSKARBDKCGGSDNSCKTVAETFTWYVYGVWYVIRIAPAGANNIDVR  
QHSFSGETDDNDLALSSSKGEFLNLNENFVIAKREIRIIGNAVVETSGSETRAVETIN  
STRIRIIEQLLQVLSVSKLNPVRYSFNIPEDKQDQYWNVSHGPMQACSRPCQGER  
KRKLVGTRRESDQVLDSDQCDRLPOGHILPECGTCDLMMHVASRSPQAGGLER  
TLDIYCAKYSRLDGTKEKVDGFCSSHPKSPNREKSGECNTGMRYSAWTEGSKSCD  
GTRORRALICVNRVNDYLDKSCITHOEKVYIORSSEFPDQWKSQWMSQICVITCKGKH  
KRNQVWQGFEDLNDRCMPETKPTSMOTCOQPEASWQAGRGCCSVTCCGQGYLR  
AVKCIITWMSVVDNDNCNATPPTDIDELDSCHPPEAPETRSTYSAPETQWRK  
GSMTPSANTGKGTIRMYVYSCRDENGSVDESACALIPRVAARECSVTPCGQWALD  
WSSSVYCGGGRATROYMCVNYSDHVLDRECODDIIPETDQCSASPPORTPDENGL  
AHPFQNEBDRPSPASPTNHLVGGNOMRTGPMVAGCSPGACSSRAGSRRVYVCOGMSL  
ANDCVERIKRDEDRACSSGRCPOWAYGNMGECKIKLGGGIRPLVYVCOASNGRPEDL  
SCETLDRPPREDCNTHACPHDAWSTGPNSSCSVSCGRRHKORANYCMAKDSHLES  
DCKHLAKRPHGRKRCGRCPKWKAGMWSQCSVSGRQVQDRHVGCOIGTHKLAETLE

CNPTRESEBDCGPRCPLTYTWAEEMOECTKTCGEGSRHYRKYVVDNDKKNFVHGAR  
CDVSKRPVDRSSCSLQPEYKYWITGEMSEVPSWEL"  
BASE COUNT 1471 a 1112 c 1362 g 1194 t  
ORIGIN  
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..

Initial Score = 5 Optimized Score = 9 Significance = -0.40  
Residue Identity = 19% Matches = 12 Mismatches = 39  
Gaps = 10 Conservative Substitutions = 0  
Translation Frame = 1  
OLRANRQPLRLRNAGHIEPVXWGLEFYHTTYYGKTRRXKGTQKQTRPHLHAKORPERALNRAKCAKXHLRQK  
10 20 30 40 50 60 70  
XAAQXROEENQSKKMGKDKXRGWXRSSIKORLSNRGIFCLMXXDGNKREDRQKDFEILSTVCRSLGGGR  
80 90 100 110 120 130 140  
QONGFLPWRKRPSTLYFNENVNSLYLXRPKYKFNXXCYCELNDCSXXTGMAFHIFGCSDNIKLLPYAAAE  
150 160 170 180 190 200 210  
EGRMWPXSYCCSLNKKTYLQSSRQXXYLRRPGTGNHLXSLKLFYXXRXWIEYSFYDRPYAGPCVXHA5XX  
220 230 240 250 260 270 280  
OQOMXRRSXSESPACHG5WTELLHDPRLDVKYXSKIYHXVFRNMLMRYAPATYIQTLPRACPPYARNPLDRE  
290 300 310 320 330 340 350 360  
XTPXXIDFWTRFSGVPRFYDAVOTALVQXROWSTQRLPDSNHTLGRWDVYRAMKALQWILCSORNGCPDRMI  
370 380 390 400 410 420 430  
LKLKESLWMLLNMWRGNHONSHRSVQDGTQTKKWKILCFTXNLEIYVLDHGANSQAEARLPRXVTSLXREAF  
440 450 460 470 480 490 500  
XHORASQCALGPHXIQIOWNSDEGRVQVVLQSGREHSLLSASRQSDRMNSLMPGKXKXVLCRPLPASWMSRCEK  
510 520 530 540 550 560 570  
LKSPEFXMKGLMWRKFEKMONSGRNITYSTPLMDIQCGPNSWCYQXCAAQDLEKGNRRKXOLDLSPFKQXHXIL  
580 590 600 610 620 630 640  
AKKRLCCHNGKGN5HECCGGRVOWRDCRRKXILNRSXKARTFASGFVAGVQVORPCTLEFQY5NXXHTSA  
650 660 670 680 690 700 710 720  
VLEEG5MAMASMOXTPRGTETKTCIHDGIXSAYCFXSKMRAADARVHTYXLYMRYLREVAACQOQEMXKCP  
730 740 750 760 770 780 790  
VWLGRLPHIGLILCOIXQAGMEDXGXWFLQOQSPQTKQFQXKMLRGXHXHGLALFLCIDXMKFKLKHMDREBEGK  
800 810 820 830 840 850 860  
YLQIYRXXCTGXQOQNTPSSE5H5SEVQXVPL5TWEIWLVRVLGHLMKRKAAPRGLV5VWMA5K5IXXVWXXP  
870 880 890 900 910 920 930  
DQANIYADLSAAGMSTLADGSLGTVOCHLMTGTIPRAKSS5MHNHMDLVVSGRQXHLXCSMNXTNXXYPLGIXITIMS  
940 950 960 970 980 990 1000  
1010 1020 1030 1040 1050 1060 1070

50 X  
CMTPEPC  
SPSPCGNEEKNHIOCTKRNPAIVLVLDPLMS-HLMEYRPEIDRQLPRXEMLCSXBRCLCYRPAVTSKGRMF-  
1010 1020 1030 1040 1050 1060 1070  
X  
GSMGMSRPGTSGRGGGICATAIKRCNRPREKNGKCYC-----VGRMRKFKS  
10 20 30 40  
PLPKXGLSSPEROPRPHRCARMKRVENWPLGSMFYULCMRITRAPACCSMG5XKWIHRKRLCSGEMKTXKXAKSLX  
1150 1160 1170 1180 1190 1200 1210 1220  
IRPLSSVGLMQLGRVHXAVWMMRHKNRKNTGGISAVQRTVSRFELXNSXTSRSXAVVHTCLSTRROMEMVPLF  
1230 1240 1250 1260 1270 1280 1290  
LVFCLMWSRAXXTKCLLHGRKMRKPRFKKLLXLABGAXYTAQAKVPRRKMPOBESWRLE5VLCLVIMPRRTYAEAGC

1300 1310 1320 1330 1340 1350 1360  
 LSDNNTONSQRDVPQIPIHGVGTRRLRPRPIVPLHLEGMRARHODLRRRLQVPOGGVCGXOQKRGAWGTL  
 1370 1380 1390 1400 1410 1420 1430  
 XREDAAGGPKLXKATLRLVLDLRRMVRGVLTGVTIVSSAMAXEMQRDECRNRVGLRLRPARPAAVNC  
 1440 1450 1460 1470 1480 1490 1500  
 VFSDKSSAFPSXNSLVAKRERCKRWSLSPFRSTPCDSPXCPGIGIKKLLGLQDTHMLXGAXRQLSIRKXCS  
 1510 1520 1530 1540 1550 1560 1570 1580  
 ARDTRRDGNDKLAHAYXREQNLPHSSCGHAEMKSSVTRPDPFSREAKNLDXYVRRPRLKMGAEIFGFKTLX  
 1590 1600 1610 1620 1630 1640 1650  
 GANKPLSXTRCVVXQOTASEENNPLXPKYKLYHPFYLYPHNAVYLDPEFXMXXTSSNH  
 1660 1670 1680 1690 1700 1710

6. US-09-972-467-2 (589-642)  
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139  
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000  
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.  
 ACCESSION AB037733  
 VERSION AB037733.1 GI:7242978  
 KEYWORDS  
 SOURCE Homo sapiens brain cDNA to mRNA, clone\_lib:pbluescriptII SK plus  
 clone:fh11767.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (sites)  
 Nagase, T., Kikuno, R., Ishikawa, K. I., Hikosawa, M. and Ohara, O.  
 Prediction of the coding sequences of unidentified human genes.  
 XVI. The complete sequences of 150 new cDNA clones from brain which  
 code for large proteins in vitro  
 DNA Res. 7 (1), 65-73 (2000)  
 2 (bases 1 to 5139)  
 Ohara, O., Nagase, T. and Kikuno, R.  
 Direct Submission  
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,  
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba  
 292-0812, Japan (E-mail:cdna1info@kazusa.or.jp,  
 url:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,  
 Fax:+81-438-52-3914)

FEATURES  
 source  
 1..5139  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="fh11767"  
 /tissue\_type="brain"  
 /clone\_lib="pbluescriptII SK plus"  
 1..4417  
 /gene="KIAA1312"  
 <1..4417  
 /gene="KIAA1312"  
 /note="Start codon is not identified."  
 /codon\_start=2  
 /product="KIAA1312 protein"  
 /protein\_id="BAA92550.1"  
 /db\_xref="GI:7242979"  
 /translation="NSEHTAVISLCSGMIGTFRSHDGDYFTEPLQSMDEDEDEEONK  
 PHIIYRRAQPREPSTGRKADDTSEHKRHSKDKKTRAKRWERINLAGVVALNSG  
 LAIEAFSAVGNKTDNTRREKTRHRTKRFLLSYPRFVAVLVADNRMVSYHGENLQHYIL  
 TMSIVASIVKDPSTGNLINTIYVNLIVIHNEODGSPISFNAQITLTKRFGOMSKNS  
 RGSIHNDVAVILTRDIDGRANDKCDTGLAELGTTCDYTRCSLSEDSGLSTATTIAH  
 ELGAVFNMHDDNCKKKEGVKSPQHVMAPTLNFYTNMWSKCSRYTIFLDTGIG  
 ECLLNEPSRPLPVOLPGLIYVNVKCELIIFGSGVCPYMMQCRILMNNVNGVH  
 KGCRRQHTPMADTECEPDKKCYGFCVAKEMDVAIVTDSGMSMSPTGCRSTCGGI  
 KTAIRECNRPEKNGKCYCVRMKKFCSCNTEPCIKQKRDPRDSCAHFQCKHNNING

Initial Score = 4 Optimized Score = 10 Significance = -0.45  
 Residue Identity = 18 Matches = 13 Mismatches = 37  
 Gaps = 20 Conservative Substitutions = 0  
 Translation Frame = 3  
 BASE COUNT 1471 a 1112 c 1362 g 1194 t  
 ORIGIN  
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..

PSSTPSSASAOECSVAHSGIMMGIILLNHYSLMNNKMKRNTPTSTIGAAPPRESPOGEMHVTPTONTKI  
 10 20 30 40 50 60 70  
 GGIKRAIRRCNRPREKNGSK-----YGVRRMKFKSCNTERC \* X  
 20 30 40 50 60 70  
 GIVKTR---RRKPEQENGEKGLTWLVYQHXHTAAXQORHFLMVTIRFTQKRGPTREGONVFFYRHLKXSM  
 80 90 100 110 120 130  
 WMQTLWEFHTMKTENTLPELXKCOLXPLSITQVLEIYLLILLLKTKLFTMNRMGRLVTLMLRQHXTPFASG  
 150 160 170 180 190 200 210  
 SIRTPVQVESIMILLFSKQDIRISAEITLTVIPXAWLNWEPVPIEAVLIVKYVQDLRSPMAMCLTC  
 220 230 240 250 260 270 280  
 IAAITLTVNKKKRELRVPSMSWLQHXSTPTPGGOSVVENISLSRFTLVMAVCLTLNPPDPTLGLSNCOASF  
 290 300 310 320 330 340 350  
 TTXINNVMXFLDQVLRCAHIXCSADGSGAITSMEYTKAAGLSTHRGPMGRSASLSTPAMDFVPPKRWMSPX  
 360 370 380 390 400 410 420  
 QMDPEVGVLEPAPAEVHEGASKQPFESATDONKQMVENYXADVEXNLSPTARRSHVSSRSTSMNSVLTLT  
 430 440 450 460 470 480 490 500  
 GSILFSTVGFPRCAGSLNTVDEYXKRTGASCSAEWQGTQPTLISFETEXMELLVARTOMISVSAFAKDLDAI  
 510 520 530 540 550 560 570  
 MFXTKQPEINWGVVALIILNAKQOEHLIQYIMWTLIMSEFQVLPILPIMGSGSVVSOGKQMTTXXLXQAVK  
 580 590 600 610 620 630 640  
 VNSCAMEPILLSQWPKGKRALGMLMXSTVGPRLPKKELTQDIALSKNFCFRRCWESCTTPMYAIIISIFQDKI  
 650 660 670 680 690 700 710  
 NLSSTGTVMGHGKIAAVNPAKNGNENLFAFGNLSLFLFKADIGCSLDTLLNPPVQVTVXGGMLDAGVN  
 720 730 740 750 760 770 780  
 VVPSAVAVTVAHTSTVIVPIAGMGLRLRYMVAALPNDQATVKNMAQGNVTRVAGAILPGLINOKAVTVGPR  
 790 800 810 820 830 840 850 860  
 GEGLFVSIPEMNYWMTANAHLKRRLLPFRGAVSVLHSGNLETTGQSAMSPEYKIGISTARSGSLVKKIDXMTIEC  
 870 880 890 900 910 920 930  
 VTLRPSQHCRLVSSRNVAHPGRVPGDSAVSLVDRDPTSKXEQKASLSGLICQMKMTVWQDLDLIPRTVNY

940 950 960 970 980 990 1000  
 HHVILPDLPRKREAHPTVHQPSSDGLGPHADPLVGGKFGKPTSAEMKMLLTVPLDCLDQWQRKNV  
 1010 1020 1030 1040 1050 1060 1070  
 LXHPVGNRPMGALALXPKVKGQPGKCVSTTWTXSIGVSVFRHISQKLTTRVPCHHALKGPOTVAXLS  
 1080 1090 1100 1110 1120 1130 1140  
 TPSEMTIYVGAAPAAPMGSVETSGELARGEHVVPVVLADPSSGVLIVYMKMDTPQTTVMREXNLSKPEV  
 1150 1160 1170 1180 1190 1200 1210 1220  
 NPALVLSGLMATGSEALS CVVEAXXEDPMSVSGPTVNGFOIXAVKFLINLPYVSVYTHMLVHTTLLHGVLALG  
 1230 1240 1250 1260 1270 1280 1290  
 ARVLSLYVEGINNEMFTANQKMEAIYKVTITVSTWLSHMGTEESAEEEDAPRNGKIALGVSAALCPVAEAYSRGMW  
 1300 1310 1320 1330 1340 1350 1360  
 AVRESEHTKXPERPSATHTPDRSRNATKAKHGVSPSTLGGQRNGKNAPRPAKAPGATARMCVMMTTKTRCMGHA  
 1370 1380 1390 1400 1410 1420 1430  
 VTYASGRTWYKAVVNCNPASMSGQENQGRYRPGNCHNRQLSHGIRVAEGXVEGXVQECGRLEATRPICHCEL  
 1440 1450 1460 1470 1480 1490 1500  
 CVFXQVLSFPKLEFLVCKAGEEMKEMVSKSLQVYILKFTLMSYWHKKEIITGAANSKXAVRCIKTVKYYKKILX  
 1510 1520 1530 1540 1550 1560 1570 1580  
 CQGNKEREMQLXOTSTCCVGAESLSLQIWPICRNWVYXRQYTFPKRKGKSGLVCEISXLENGGXNPFMFXNIV  
 1590 1600 1610 1620 1630 1640 1650  
 RQKQTPPMNQMCACALTNFSFRGRKXFSIISEVSOVPEFHISCAQCISRPFLMWINOXXS  
 1660 1670 1680 1690 1700 1710